Research Paper

Development and external validation of a dynamic prognostic nomogram for primary extremity soft tissue sarcoma survivors

Dario Callegaro[a], Rosalba Miceli[b], Sylvie Bonvalot[c], Peter C. Ferguson[d], Dirk C. Strauss[e], Veroniek V.M. van Praag[f], Antonin Levy[g], Anthony M. Griffin[d], Andrew J. Hayes[e], Silvia Stacchiotti[h], Cecile Le Pèchoux[g], Myles J. Smith[e], Marco Fiore[a], Angelo Paolo Dei Tos[i], Henry G. Smith[e], Charles Catton[j], Joanna Szkandera[k], Andreas Leithner[l], Michiel A.J. van de Sande[f], Paolo G. Casali[h,m], Jay S. Wunder[d], Alessandro Gronchi[a,*]

[a] Department of Surgery, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan 20133 , Italy
[b] Department of Clinical Epidemiology and Trial Organization, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
[c] Department of Surgery, Institut Curie, PSL Research University, Paris, France
[d] University Musculoskeletal Oncology Unit, Mount Sinai Hospital, Department of Surgical Oncology, Princess Margaret Cancer Center, Department of Surgery, University of Toronto, Toronto, ON, Canada
[e] Department of Surgery, Royal Marsden Hospital NHS Foundation Trust, London, UK
[f] Department of Orthopaedic Surgery, Leiden University Medical Center, Leiden, The Netherlands
[g] Department of Radiation Oncology, Gustave Roussy Cancer Campus, Paris Saclay University, Villejuif, France
[h] Department of Cancer Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
[i] Department of Pathology, General Hospital of Treviso, Treviso, Italy
[j] Department of Radiation Oncology, Mount Sinai and Princess Margaret Hospitals, University of Toronto, Toronto, ON, Canada
[k] Division of Clinical Oncology, Medical University of Graz, Graz, Austria
[l] Department of Orthopaedics and Trauma, Medical University of Graz, Graz, Austria
[m] Oncology and Hemato-Oncology Department, University of Milan, Milan, Italy

A R T I C L E   I N F O

A B S T R A C T

Background: Prognostic nomograms for patients with extremity soft tissue sarcoma (eSTS) typically predict survival or the occurrence of local recurrence or distant metastasis at time of surgery. Our aim was to develop and externally validate a dynamic prognostic nomogram for overall survival in eSTS survivors for use during follow-up.

Methods: All primary eSTS patients operated with curative intent between 1994 and 2013 at three European and one Canadian sarcoma centers formed the development cohort. Patients with Fédération Française des Centres de Lutte Contre le Cancer (FNCLCC) grade II and grade III eSTS operated between 2000 and 2016 at seven other European reference centers formed the external validation cohort. We used a landmark analysis approach and a multivariable Cox model to create a dynamic nomogram; the prediction window was fixed at five years. A backward procedure based on the Akaike Information Criterion was adopted for variable selection. We tested the nomogram performance in terms of calibration and discrimination.

Findings: The development and validation cohorts included 3740 and 893 patients, respectively. The variables selected applying the backward procedure were patient's age, tumor size and its interaction with landmark time, tumor FNCLCC grade and its interaction with landmark time, histology, and both local recurrence and distant metastasis (as first event) indicator variables. The nomogram showed good calibration and discrimination. Harrell C indexes at different landmark times were between 0.776 (0.761−0.790) and 0.845 (0.823−0.862) in the development series and between 0.675 (0.643−0.704) and 0.810 (0.775−0.844) in the validation series.

Interpretation: A new dynamic nomogram is available to predict 5-year overall survival at different times during the first three years of follow-up in patients operated for primary eSTS. This nomogram allows physicians

to update the individual survival prediction during follow-up on the basis of baseline variables, time elapsed from surgery and first-event history.

Research in context

Evidence before this study

We searched PubMed for studies published before Nov 1, 2018, that investigated the use of nomograms for predicting prognosis of patients with extremity soft tissue sarcoma. We used the search terms 'nomogram', 'prediction', 'sarcoma', and 'extremity'. Available prognostic nomograms for patients with extremity soft tissue sarcoma predicted survival or the occurrence of local recurrence or distant metastasis at time of surgery. After surgery, the individual risk of dying shows a non-linear evolution which is determined by the time elapsed from surgery, the event history (occurrence of local recurrence or distant metastasis vs. no events), and by the time-dependent effect of baseline covariates (tumor and patient characteristics). At the time of our search, there were no nomograms available for use during follow-up. The only available dynamic prognostic model was limited to patients with high-grade tumors, adopted a suboptimal histological classification and was not externally validated.

Added value of this study.

With this study, we created and externally validated a dynamic prognostic nomogram for overall survival to obtain a longitudinal prognostic estimation for patients with primary extremity soft tissue sarcoma. In particular, this nomogram allows individual prognosis prediction at baseline and at different time points up to three years after surgery. The main strengths of this nomogram are the use of easily definable covariates, which were all strongly associated with survival, and the successful external validation. This nomogram has been incorporated in the app 'Sarculator' for smartphones and tablets.

Implications of all the available evidence.

With this new nomogram, physicians are now able to inform patients of their residual risk during follow-up. Psychologically, this may help patients coping with the fear of cancer recurrence. The longitudinal risk estimation can also aid personalization of follow-up policies based on the individualized residual risk. Finally, with this instrument physicians are able to quantify the prognostic impact of local relapse and distant metastasis in a specific clinical scenario, helping decision-making in patients who recur.

---

1. Introduction

Extremity soft tissue sarcomas (eSTS) are characterized by a broad histological diversity which results in a substantial variation in clinical course [1].

Prognosis estimation for primary eSTS patients has progressed in recent years with the creation of dedicated prognostic nomograms. With these tools, physicians are able to compute a personalized prognosis on the basis of baseline clinical variables such as patient's age, tumor size, grade and histology [2,3].

Prognostic nomograms for eSTS patients typically predict survival or the occurrence of an event, namely local recurrence (LR) or distant

metastasis (DM), at time of primary treatment. These static predictions may aid the therapeutic decision making process, for example identifying patients at higher risk of death or disease recurrence who may benefit from a combination therapy [2,4–8]. However, during follow-up (FU) patient prognosis will change depending on the time elapsed from primary surgery, the event history (LR vs. DM vs. no events), and related to the possible time-dependent effect of covariates. As such, "static" nomograms do not provide accurate predictions if used at a particular time point during FU.

In the past few years there have been studies exploring survival-specific prognostic information in patients with soft tissue and bone sarcoma showing that the effect of prognostic variables changes as patient's survival time increases and that, without events, the likelihood of surviving improves as time goes by after treatment [9–13].

Determining the patient's risk of death or tumor relapse at a certain time-point given a set of covariates and the event-history until that moment goes under the broad methodological framework of "dynamic prediction". One way to achieve dynamic predictions is to perform a landmark analysis [14]. This technique involves taking a 'snapshot' at a given time, with the creation of a "landmark dataset" including only the patients at risk at that time point. All the landmark datasets are stacked in a "super dataset" on which a single regression analysis (e.g., "Cox supermodel") can be performed. The first dynamic nomogram based on landmark analysis was developed for breast cancer patients in 2015 proving the feasibility of this approach [15].

The aim of the present study was to develop and externally validate a dynamic prognostic nomogram to predict overall survival (OS) in primary eSTS survivors at different times during follow-up.

2. Methods

2.1. Study design

This is a retrospective analysis. All consecutive adult (>18 years) patients with primary (non-recurrent, non-metastatic) eSTS surgically treated at Fondazione IRCCS Istituto Nazionale dei Tumori (Milan, Italy), Institut Gustave Roussy (Villejuif, France), Mount Sinai Hospital (Toronto, Canada), and at the Royal Marsden Hospital (London, UK) from 1994 to 2013 were merged, forming the development cohort. On the Milan series, we developed two static nomograms for OS and DM in 2016 [2]. Patients with the same characteristics operated on between 2000 and 2016 at 7 other European referral centers comprised the validation cohort. The full list of participating centers is available in Supplementary material.

Extremity STS were defined as tumors arising between the shoulder girdle and the hand (upper extremity) and between the pelvic girdle (excluding endopelvic tumours) and the foot (lower extremity). Patients with well-differentiated liposarcoma, dermatofibrosarcoma protuberans, desmoid-type fibromatosis, Ewing sarcoma and alveolar or embryonal rhabdomyosarcoma were excluded.

Tumor margins were classified as microscopically positive (R1) or negative (R0). Patients who underwent macroscopically incomplete (R2) resections were excluded. The FNCLCC criteria (grades I, II, and III) were applied for tumor grading [16].

Histologically, tumors were classified according to the WHO's criteria and patients were grouped into nine categories [1].

Radiotherapy (RTx) and/or chemotherapy (CTx) were used according to multidisciplinary guidance or as part of clinical trials. The follow-up strategy consisted of clinical examination and imaging

studies, every 4 months for the first 2 years, every 6 months until the fifth year and yearly thereafter.

2.2. Ethics

The study was approved by the Institutional Ethics Committee at Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy (Study Protocol: INT 98/15). This research did not require specific written consent of the patients.

2.3. Statistical analysis

The dynamic nomogram was developed using a landmark analysis approach and a multivariable Cox model [14]. A backward procedure based on the Akaike Information Criterion was adopted for variable selection [17]. Nomogram performance was tested in terms of calibration and discrimination. The analyses were performed with SAS

(Cary, NC, USA) and R software [18]. Statistical methods are detailed in the Supplementary material.

3. Results

The multicenter series from which we extracted the development cohort totaled 3752 patients; 12 patients were excluded because survival time was missing, thus the development cohort included 3740 patients. The validation cohort consisted of 893 patients with grade II (12.8%) and grade III (87.2%) eSTS. Clinicopathological characteristics of the two cohorts are listed in Table 1. The median follow-up was (interquartile [IQ] range) 79 months (44–119 months) for the development cohort and 71 months (43–108 months) for the validation cohort. In the development and validation cohorts, respectively, 1003 and 367 patients died, including 259 and 55 patients who died without evidence of LR or DM; 263 and 99 patients developed LR; 951 and 482 developed DM as first event; 60 and 180 developed

Table 1
Demographic, clinical, and pathological characteristics of the development and validation cohorts.

	Development Cohort		Validation Cohort		
	No.	%	No	%	p value*
Total	3740		893		
Patient's age, years					<0.0001
Median	56		62		
IQR	42–69		49–73		
Sex					0.8142
Female	1692	45.2	357	44.7	
Male	2048	54.8	441	55.3	
N/A	0		95		
Depth					0.3274
Deep	2615	69.9	636	71.6	
Superficial	1125	30.1	252	28.4	
N/A	0		5		
Tumor size, cm					<0.0001
Median	6.85		8.0		
IQR	4–11		5–12		
Histological subtype					<0.0001
LMS	478	12.8	43	4.8	
DD/pleom lipo	197	5.3	37	4.1	
Myxoid lipo	510	13.6	2	0.2	
MPNST	201	5.4	78	8.7	
Myxofibro	446	11.9	189	21.2	
Synovial sarcoma	288	7.7	107	12.0	
UPS	796	21.3	357	40.0	
Vascular sarcoma	73	2.0	16	1.8	
Other	751	20.1	64	7.2	
FNCLCC grade					<0.0001
I	665	17.8	0	0	
II	1233	33.0	114	12.8	
III	1842	49.3	779	87.2	
Surgical margins					<0.0001
R0	3157	84.4	426	61.6	
R1	583	15.6	266	38.4	
NA	0		201		
Chemotherapy					<0.0001
done	641	17.1	46	6.4	
not done	3099	82.9	671	93.6	
N/A	0		176		
Radiotherapy					<0.0001
done	2232	59.7	632	76.3	
not done	1508	40.3	196	22.7	
N/A	0		65		

Abbreviations: IQR, interquartile range; R0, complete resection with microscopically negative margins; R1, complete resection with microscopically positive margins; LMS, leiomyosarcoma; DD/pleom lipo, dedifferentiated/pleomorphic liposarcoma; Myxoid lipo, myxoid liposarcoma; MPNST, malignant peripheral nerve sheath tumor; Myxofibro, myxofibrosarcoma; UPS, undifferentiated pleomorphic sarcoma; FNCLCC, French National Federation of the Centers for the Fight Against Cancer. * p values at Wilcoxon Mann–Whitney test or exact Fisher test, depending on whether the variable is continuous or categorical, for testing the difference between development cohort and validation cohort.

synchronous LR and DM. In the development cohort, 5-year OS was 76.0% (74.6–77.5%) and 10-year OS was 66.3% (64.3–68.2%). In the validation cohort 5- and 10-year OS was 59.5% (56.0–63.1%) and 48.0% (43.8–52.6%), respectively. Supplementary Fig. 1 shows the OS curves (panel A; p value at log-rank test<0.00001) and the crude cumulative incidence curves of LR and DM (panel B; p value at Gray test: 0.0002 for LR and <0.00001 for DM) in the development and validation cohorts. Supplementary Figs. 2 and 3 show the number of patients in the development and validation cohorts at different T_{LM}s according to LR or DM status. In the development series, OS and CCI of LR and DM were not significantly different between patients operated before or after 2005 (Supplementary Fig. 4).

3.1. Dynamic OS nomogram

In the multivariable Cox landmark OS supermodel, after application of the backward procedure the following variables were excluded from the covariates set: tumor's depth, surgical margin status, CTx administration, RTx administration. The final supermodel included age at surgery, tumor size and its interaction with T_{LM}, grading and its interaction with T_{LM}, histology, and both LR and DM indicator variables. All selected variables had a significant effect on OS (all Wald tests $p<0.0001$; Table 2). For both tumor size and tumor grade the hazard ratios (HRs) decreased at increasing T_{LM}, indicating that the strong association with OS characterizing these variables at baseline became gradually weaker the longer that patients remained alive during follow-up.

The OS dynamic nomogram derived from the final model (Fig. 1) allows the computation of the 5-year probability of being alive or, alternatively by its subtraction from 1, the 5-year probability of dying from any cause. Prognosis can be estimated at the time of the primary surgery or at 12, 24 or 36 months after the primary surgery. The prediction window is fixed at 5 years, meaning that, for example, a prediction computed 1 year after surgery provides the probability of being alive 6 years after surgery. For ease of computation the scores related to the variable 'tumor size' vary according to 'tumor grade' and T_{LM}, but this does not imply a statistical interaction between tumor size and tumor grade, as the model only included interactions between tumor size by T_{LM}, and grading by T_{LM}. A digital version of this prediction model has been implemented in the version 2.0 of the app 'Sarculator' (www.sarculator.com). Here, prognosis can be estimated at every three-month interval from the time of the primary surgery up to three years of follow-up.

Fig. 2 is an example of how the 5-year OS dynamic prediction varies during follow-up according to an individual patient's event history (data shown in Supplementary Table 1 and 2). After surgery, in the absence of events, the individual risk of dying decreases as time goes on. This is due to the fact that the population of patients alive and without events at time x after surgery will have more favorable characteristics (absence of LR and DM) as compared to the population which, at surgery, share the same set of baseline covariates. This is because in the landmark model patients who already developed LR and/or DM or who already died before time x will no longer be included in the risk set at time x. Fig. 3 and Supplementary Tables 3 and 4 reproduce the two examples of Fig. 2 but using a different tumor size.

3.2. Nomogram calibration and discriminative ability on development and validation cohorts

Calibration plots for internal and external calibration are shown in Fig. 4. Calibration plots for development series patients operated before or after 2005 are shown in Supplementary Fig. 5.

In the development series, the Harrell C index was (95% bootstrap confidence interval) 0.776 (0.761–0.790) for predictions calculated at time of primary surgery ($T_{LM}=0$) and 0.837 (0.822–0.851), 0.845

Table 2
Final multivariable Cox landmark supermodel for overall survival.

Covariates	HR	95% CI	P
Age, years†			<0.0001
69 vs. 42	1.80	(1.58,2.05)	
Local recurrence			<0.0001
yes vs. no	5.63	(4.26,7.44)	
Distant Metastasis			<0.0001
yes vs. no	10.34	(8.74,12.23)	
Histological subtype			<0.0001
LMS vs. Myxoid lipo	1.78	(1.26,2.52)	
DD/pleom lipo vs. Myxoid lipo	1.37	(0.93,2.02)	
MPNST vs. Myxoid lipo	1.73	(1.16,2.58)	
Myxofibro vs. Myxoid lipo	1.05	(0.72,1.53)	
Synovial sarcoma vs. Myxoid lipo	2.03	(1.43,2.88)	
UPS vs. Myxoid lipo	1.18	(0.85,1.63)	
Vascular vs. Myxoid lipo	3.20	(1.85,5.53)	
Other vs. Myxoid lipo	1.48	(1.07,2.04)	
Size, cm (T_{LM}, mo)†			<0.0001
11 vs. 4 (0)	3.06	(2.53,3.70)	
11 vs. 4 (12)	2.32	(1.92,2.80)	
11 vs. 4 (24)	1.90	(1.55,2.32)	
11 vs. 4 (36)	1.65	(1.29,2.11)	
FNCLCC grade (T_{LM}, mo)			<0.0001
II vs. I (0)	2.55	(1.75, 3.73)	
II vs. I (12)	2.07	(1.42, 3.01)	
II vs. I (24)	1.63	(1.11, 2.40)	
II vs. I (36)	1.26	(0.82, 1.94)	
III vs. I (0)	4.88	(3.40,7.02)	
III vs. I (12)	2.59	(1.79,3.75)	
III vs. I (24)	1.59	(1.08, 2.33)	
III vs. I (36)	1.09	(0.72,1.67)	

In principle, the HRs could be estimated at the times from 0 to 36 months; we choose to represent in the Table only four "significant" times: 0, 12, 24 and 36 months. The HR associated to size 11 cm vs. 4 cm tumors was 3.06 at baseline and dropped down to 1.65 at three years after surgery, and the HR associated to grade III vs. grade I tumors was 4.88 at baseline and dropped down to 1.09 at three years after surgery. The occurrence of DM and LR were associated with an increased risk without a time-varying effect (the interaction time by LR or DM were not retained in the final model according to the backward selection).

† The two values are, respectively, the 3rd and 1st quartiles of the variable distribution. Abbreviations: OS, overall survival; HR, hazard ratio; 95% CI, 95% Confidence Interval; P, p value at Wald test; LMS, leiomyosarcoma; DD/pleom lipo, dedifferentiated/pleomorphic liposarcoma; Myxoid lipo, myxoid liposarcoma; MPNST, malignant peripheral nerve sheath tumor; Myxofibro, myxofibrosarcoma; UPS, undifferentiated pleomorphic sarcoma; T_{LM}, landmark time; FNCLCC, French National Federation of the Centers for the Fight Against Cancer.

(0.823–0.862) and 0.834 (0.811–0.859) for predictions calculated at 1 year, 2 years and 3 years after surgery, respectively. In the validation series, the Harrell C index was 0.675 (0.643–0.704) at $T_{LM}=0$, 0.773 (0.740–0.801) at $T_{LM}=12$ months, 0.810 (0.775–0.844) at $T_{LM}=24$ months and 0.796 (0.751–0.834) at $T_{LM}=36$ months.

4. Discussion

In this study, we developed and externally validated a dynamic prognostic nomogram to calculate the 5-year survival probability at different time points over the first three years of FU in patients operated with curative intent for a primary localized eSTS. With this new tool, the prognosis estimate can be computed at baseline (primary surgery) and at any time point up to 3 years after surgery. Also, the effect of the occurrence of LR or DM has been incorporated. Both in the development and validation series, this nomogram showed good discrimination and calibration as proof of its generalizability.

Three clinical needs prompted the development of this new nomogram.

First, during FU patients often desire an estimate of their residual survival. In a recent review of advocacy priorities within the sarcoma community, the issue of survivorship with uncertainties was emphasized [19]. Without dynamic models, to update prognosis during FU
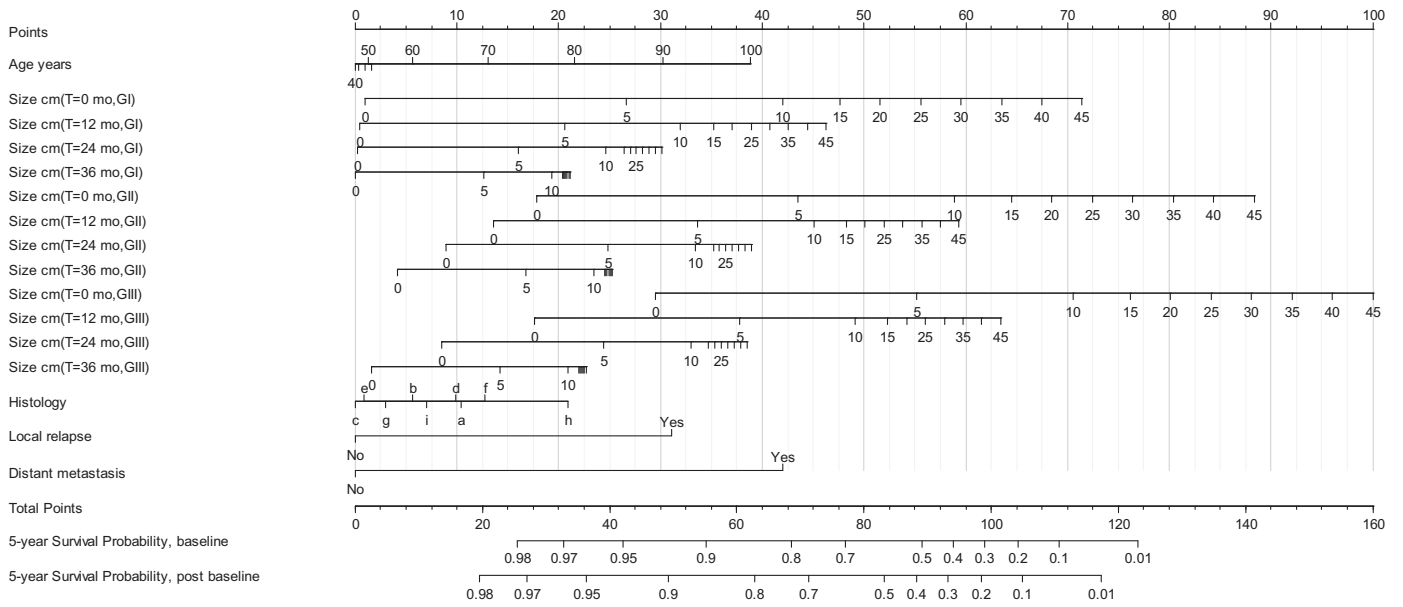
Fig. 1. Overall survival dynamic nomogram. Dynamic nomogram for 5-year overall survival (OS) in patients with primary resected eSTS. The nomogram allows the user to compute the 5-year survival probability on the basis of patient's age, tumor size, tumor grade, histology, occurrence of distant metastasis or occurrence of local recurrence as first events. The prediction can be calculated at baseline (time 0, at surgery) or post baseline (at 1 year, 2 years or 3 years after surgery). The prediction window is fixed at 5 years. Importantly, the choice of the proper size axis depends upon tumor grade and time at which the prediction is computed. For example, (i) to predict at time 0 (surgery) the 5-year OS of a 60 year-old patient with a 5 cm G3 MPNST, the user should locate patient's tumor size on the axis 'Size cm ($T = $=0 mo, GIII)' and draw a line straight upward to the "Points" axis to determine the score associated with that size and grade (55 points). Then repeat the process for patient's age at surgery (60 years = 6 points), histologic subtype (10 points), and sum the scores achieved for each covariate (sum=71 points). Locate this sum on the "Total Points" axis. Draw a line straight down to the "5-year survival probability, baseline" axis to find the predicted probability (about 78%). (ii) to predict 5-year OS of a patient with the same clinical characteristics as above but who had been operated two years before and has not had any events during the first two years of follow-up, the user should locate patient's tumor size on the axis 'Size cm ($T = $=24 mo, GIII)' and draw a line straight upward to the "Points" axis to determine the score associated with that size and grade (24 points). Then repeat the process for patient's age at surgery (60 years = 6 points), histologic subtype (10 points), local relapse (0 points), distant metastasis (0 points) and sum the scores achieved for each covariate (sum=40 points). Locate this sum on the "Total Points" axis. Draw a line straight down to the "5-year survival probability, post baseline" axis to find the predicted probability (about 94%). Histology abbreviations: (a), leiomyosarcoma; (b), pleomorphic/DD liposarcoma; (c), myxoid liposarcoma; (d), MPNST; (e), myxofibrosarcoma; (f), other; (g), synovial sarcoma; (h), UPS; (e), vascular sarcoma.
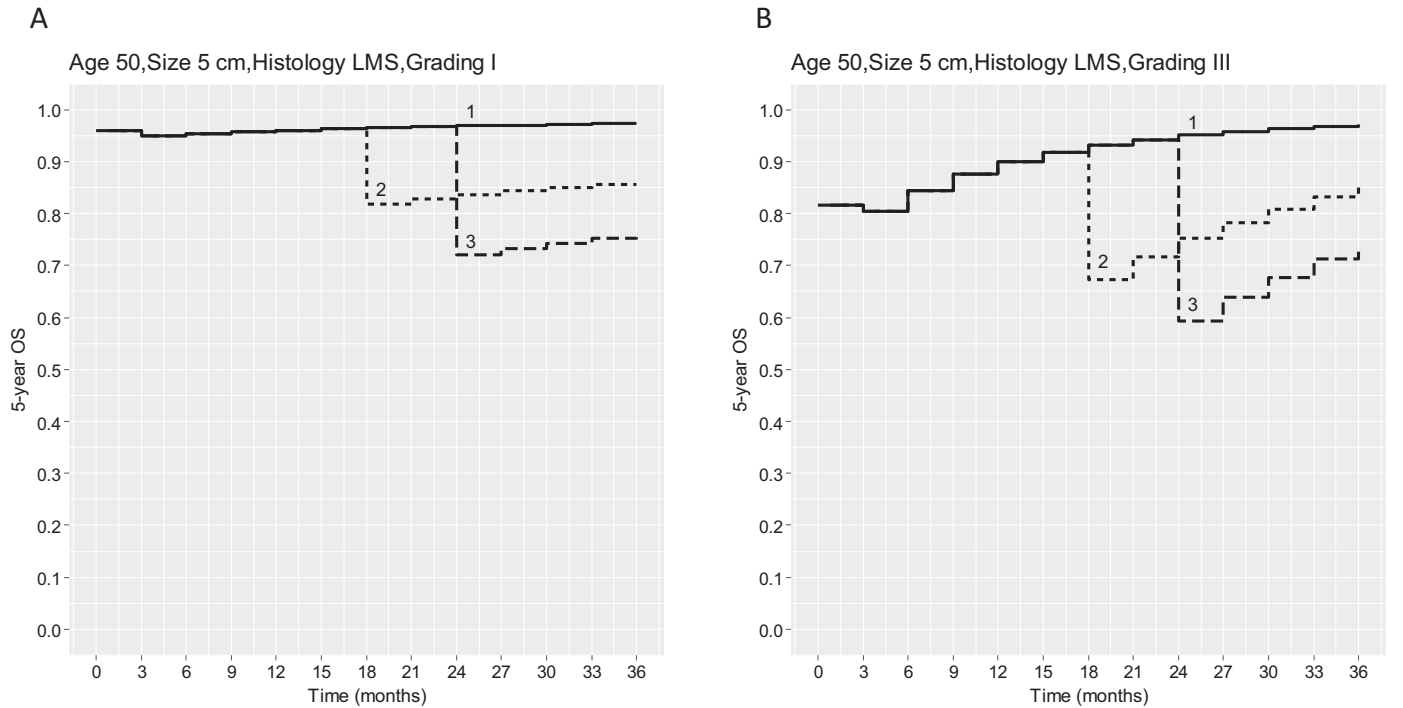
A

B

Fig. 2. Change in 5-year Overall Survival probability over time in different hypothetical clinical scenarios for individual patients. Five-year OS nomogram-predicted probabilities plotted against the time at which the prediction is computed. Line 1: uneventful follow-up. Line 2: occurrence of LR at 18 months. Line 3: occurrence of DM at 18 months. Panel A: 50-year-old patient operated for a 5 cm, GI, LMS. The predicted 5-year OS probability rises from 95.9% at baseline to 96.9% after two years of FU without events and to 97.4% after 3 years of FU without events. On the contrary, the prognosis of patients who develop LR or DM during FU drops significantly as compared to the baseline prediction, i.e. from 96.0% at 12 months to 72.1% at 24 months if the patient developed DM. At further uneventful follow-up, the 5-year survival probability will start to rise again. Panel B: 50-year-old patient operated for a 5 cm, GIII, LMS. The higher grade contributed to a global lowering of 5-year OS predictions. Lines 1−3 refer to the same event occurrences as in Panel A above. Abbreviations: OS, Overall Survival; LMS, leiomyosarcoma, LR: local relapse; DM, distant metastasis.
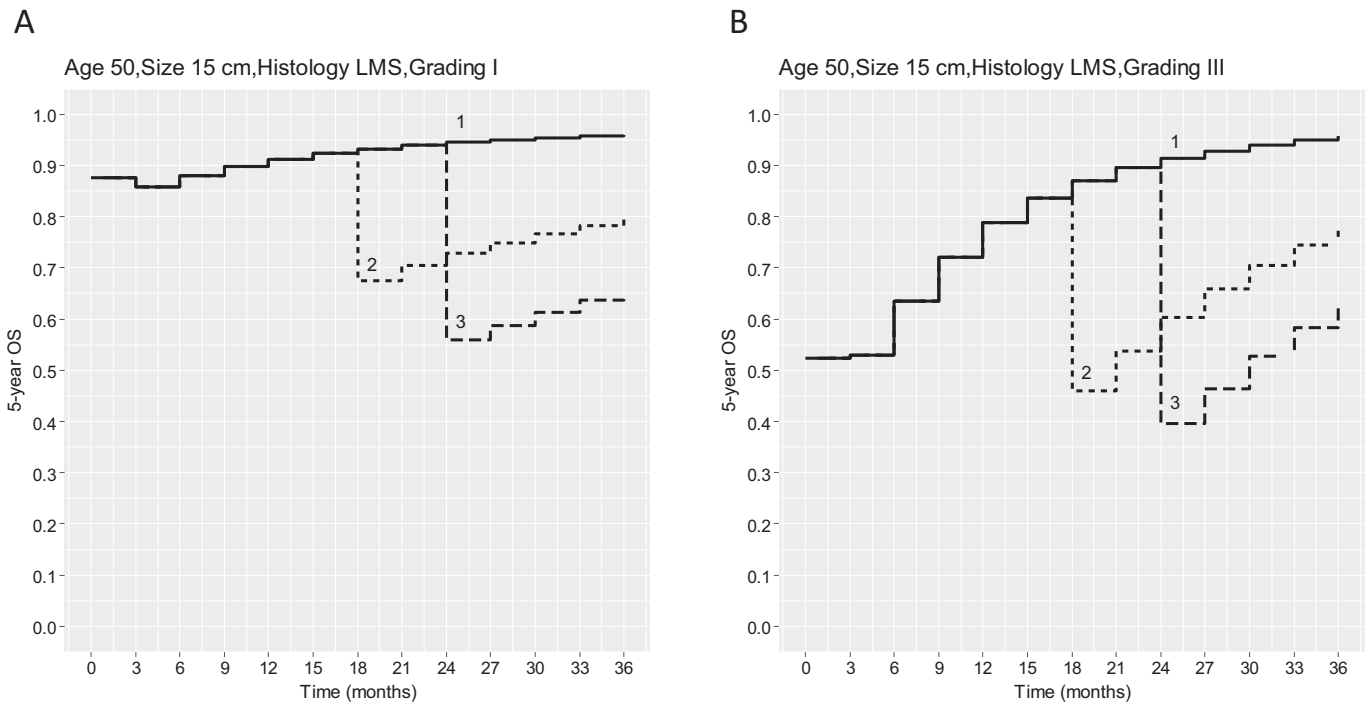
A

B

Fig. 3. Change in 5-year Overall Survival probability over time in different hypothetical clinical scenario. Five-year OS nomogram-predicted probabilities plotted against the time at which the prediction is computed. Line 1: uneventful follow-up. Line 2: occurrence of LR at 18 months. Line 3: occurrence of DM at 18 months. Panel A: 50 year-old patient operated for a 15 cm, GI, LMS. The predicted 5-year OS probability rises from 87.6% at baseline to 94.5% after two years of FU without events and to 96.1% after 3 years of FU without events. On the contrary, the prognosis of patients who develop LR or DM during FU drops significantly as compared to the baseline prediction, i.e., from 91.2% at 12 months to 55.8% at 24 months if the patient developed DM. At further uneventful follow-up, the 5-year survival probability will start to rise again. Panel B: 50 year-old patient operated for a 15 cm, GIII, LMS. Line 1: uneventful follow-up. Line 2: occurrence of LR at 18 months. Line 3: occurrence of DM at 18 months. Abbreviations: OS, Overall Survival; LMS, leiomyosarcoma, LR: local relapse; DM, distant metastasis.
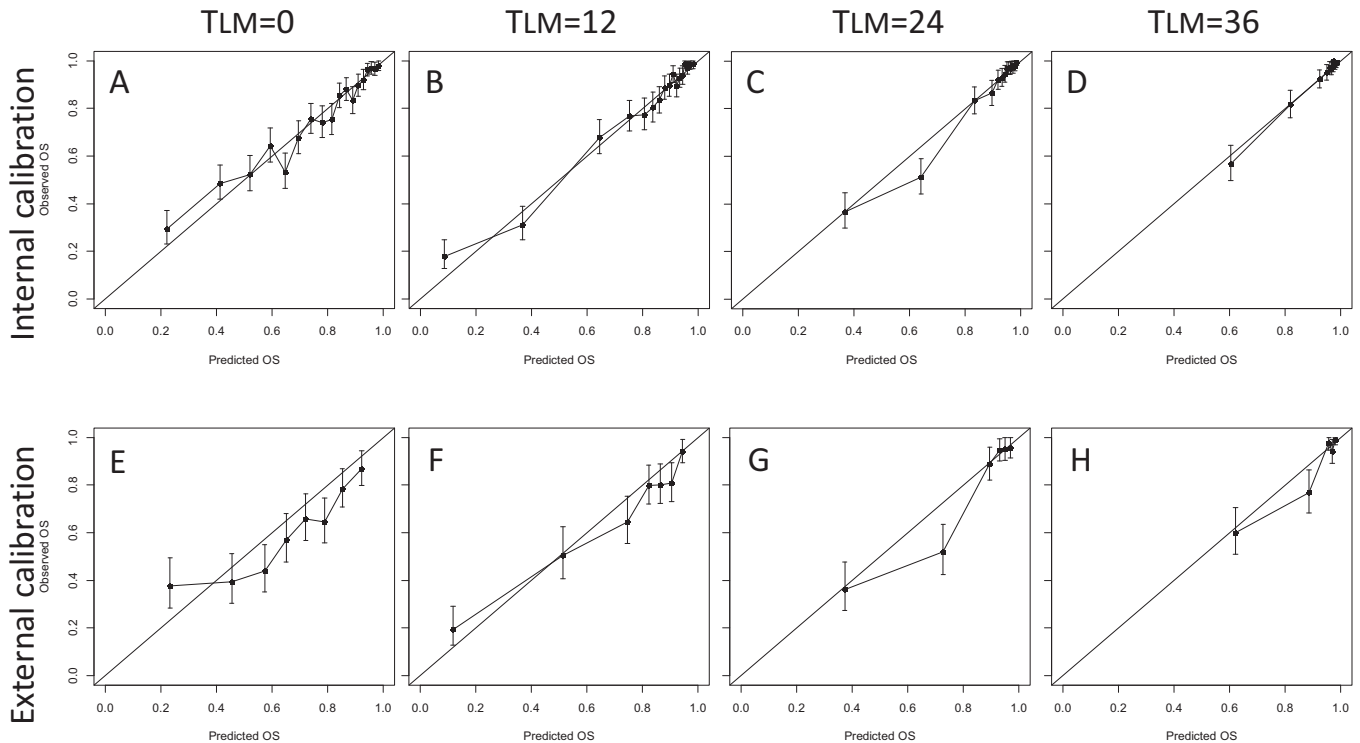

Fig. 4. Calibration plots of the dynamic Overall Survival nomogram in the development and validation series. Five-year OS at baseline (A) and (E), 12 months (B) and (F), 24 months (C) and (G) and 36 months (D) and (H). Data are from the development cohort (A)–(D) and from the validation cohort (E)–(H). The nomogram predicted probabilities were stratified in equally sized subgroups. For each subgroup, the average nomogram-predicted probability (x-axis) was plotted against the Kaplan–Meier probability observed in the same subgroup (y-axis). The 95% CIs of the Kaplan–Meier estimates are indicated with vertical lines. The continuous line indicates the reference line, where an ideal nomogram would lie. Abbreviations: TLM, landmark time; OS, overall survival.

physicians could only rely on baseline predictions and then attempt to adapt them in light of the survival curves in the specific scenario, with potentially inaccurate results. For cancer patients, a very well-known psychological adverse effect of the natural history of their disease is the so called Damocle's syndrome [20–22]. This implies that the cancer patient lives fearing relapse and cancer-related death, in spite of the fact that it will not occur in an increasingly higher proportion of patients. This certainly contributes to make the notion of cure in oncology so problematic. It is true that the residual risk of dying of disease is not nil even after many years, but also that late relapses are rare, though with marked differences from cancer to cancer [23]. Actually, patient prognosis improve with time, at a non-steady speed. Psychologically, therefore, it may be very important for the cancer patient to be updated about the evolution with time of his/her risk profile. He/she might be surprised realizing which is the residual risk. Updating the patient on a contact-by-contact basis may strengthen the information process about risks and may overcome misunderstandings related to dichotomous information (by which, say, before 5 years the risk would be high, and beyond it would be nil). This nomogram may improve patient information during follow-up in the STS field, and serve as a model in oncology.

The second need is personalizing follow-up policies. The need of improving and personalizing FU strategies was identified as a priority by advocacy groups [19]. This nomogram will allow patients to improve their risk perception during follow-up. This awareness, will lay the foundation for a shared decision making and physicians will be able to actively engage patients into choosing a follow-up strategy that could best suit their individual preferences and values on the basis of a personalized dynamic risk estimation. A shared decision making is particularly important in a scenario like this where there is no right or wrong path (i.e. intensive vs. non-intensive follow-up) and a diagnostic test can trigger patient anxiety as well as a number of other investigations [24]. Of course, there is an issue of efficacy about follow-up. By and large, we lack formal demonstrations of efficacy of follow-up as such. Indeed, several randomized trials in some common cancers were negative [25,26]. It is difficult to foresee pragmatic trials on the efficacy of follow-up in STS. Clearly, follow-up policies might be modeled. It is somewhat arbitrary to model such policies in the lack of data about the added value of an earlier diagnosis. However, if the predicted prognosis were very different at different intervals, at least there could be a rationale in narrowing follow-up windows, and vice versa. Thus, an instrument such as this nomogram might help personalize follow-up policies in STS patients, even in the lack of strong data about efficacy.

The third need is to personalize the approach to patients who develop LR/DM. Treatment decisions at relapse are often challenging. Our nomogram allows quantifying the negative prognostic implication of an oncological event in a single patient and this may aid in the decision-making process in the recurrent/metastatic setting as well as it may improve patient stratification in clinical trials in the metastatic setting.

In our new nomogram, the prediction window is fixed at 5 years while the time-point at which the prediction can be computed can vary from baseline to up to three years after surgery. Over the first three years of follow-up the risk of death among eSTS patients is substantially higher than after that time point when the risk drops significantly. This is related to the tendency of the DM curves to plateau after three years in the most common histological subtypes among eSTS [27,28]. Our nomogram covers the period of time during which updating the survival prediction is needed the most.

With the backward statistical approach, tumor depth, surgical margin status, CTx and RTx administration were excluded from the nomogram covariate set, as well as their interactions with T_{LM}. This is consistent with the findings in our two previous static nomograms for OS and DM and it is not surprising since tumor depth and margin status have been shown to have limited influence on survival in other studies while RTx administration was not associated with an improved survival in randomized controlled trials [2,29–31]. With regard to CTx, it is unlikely to be able to identify a possible effect on OS in such a diverse population which included both high risk and low risk patients. A propensity score-matched analysis recently performed on the development cohort showed a non-significant 5% survival benefit associated with CTx administration [32].

In our nomogram, the prediction is derived from the following covariates: patient age at surgery, tumor size, tumor grade, histology, occurrence of LR and occurrence of DM. Patient's age and histology are static covariates whose value and HRs do not vary with time (their interactions with T_{LM} was not retained in the model). In this nomogram we adopted the 4th Edition (2013) of the WHO histological classification to categorize patients into 9 groups. Both the granularity of the histological classification and the modeling of patient age as continuous variable are added values of this nomogram [33]. Analysis of the impact of both tumor size and grade showed that the more time that elapsed from surgery, the weaker became their association with survival (Table 2). This is consistent with other studies, in particular with what was observed by Parsons et al. in a population-based conditional survival analysis of more than 6000 primary eSTS patients treated with surgery in which the HR related to tumor size and grade decreased with increased time from diagnosis [10,34]. Finally, LR and DM are dynamic covariates, whose value (yes vs. no) may vary with time; however, their time-dependent effect was weak and their interaction with T_{LM} was not selected in the final model. The occurrence of DM is the stronger risk factor for OS with an HR of 10.34 while the HR related to the occurrence of LR is 5.63. This is likely due to the fact that in eSTS an isolated LR is unlikely to directly cause a patient's death, if we exclude the very proximal locations that may recur within the pelvis/chest. Interestingly, as noted in Figs. 2 and 3, in patients with higher risk tumors the negative prognostic impact of LR and/or DM is more pronounced.

One strength of this study is the robustness of the prediction model. By merging more than 3700 eSTS patients in a unique development cohort, we were able to capture the time-dependent effect of tumor size and grade. The long median FU of the development cohort allowed us to stretch the predictions out to 5 years, which means that a prediction computed 3 years after surgery allows estimating the probability of being alive 8 years after surgery. Of note, there was a good agreement between observed and nomogram predicted OS when splitting the development series in the two subsets of patients operated in the Nineties and more recently (calibration plots in Supplementary Fig. 5). This makes us confident in extending nomogram use to future patients treated according to the more recent policies. At external validation, although calibration for a limited subgroup of patients was not as accurate as with the internal cohort, this model still showed good discrimination with Harrell C indexes of 0.675, 0.773, 0.810 and 0.796 at the four analysed timepoints (baseline, 12, 24, 36 months). These results compare favorably with the internal validation of a dynamic prognostic nomogram predicting 5-year OS after the start of adjuvant endocrine therapy in postmenopausal, endocrine-sensitive early breast cancer patients [15]. This is true especially bearing in mind that our nomogram was validated on an independent external cohort and that the development and validation cohorts were intrinsically different in terms of baseline characteristics and outcomes (Table 1, Supplementary Fig. 1). Nonetheless, we cannot infer that this nomogram is universally applicable. The external validation has been successful on a series of patients operated among reference centers that share the same treatment guidelines and in countries with similar life expectancy. The validity of this nomogram on more diverse populations, in low-volume centers or among centers that use different treatment guidelines should be tested to further strengthen the adoption of this instrument into clinical practice.

Compared to previous studies [13], our nomogram adds significant insights since it is built on a larger cohort, is valid for patients with grade I eSTS as well as higher grade tumors, adopts a more

updated and granular histological classification and included an external validation. Moreover, we were able to select easily definable covariates which were all strongly associated with survival.

This study has limitations. First, the low number of patients experiencing concurrent LR and DM (n = =60) did not allow us to separate the OS probability estimation for this subgroup of patients and they were considered as metastatic. Second, since the nomogram model was based on the first occurrence of LR or DM, it is not able to factor in the effect of a second LR or DM on survival. In particular, the post-event survival curves reflect the outcome of both patients who will experience second events as well as those who do not. In general, this nomogram should not be used if a patient experiences other events after the first LR and/or the first DM. Finally, since the external validation was performed in a population of only grade II and III eSTS, we cannot infer whether this model is generalizable to patients with grade I tumors operated outside the centers included in the development cohort.

In conclusion, this new prognostic tool fulfills a need of the oncologist dealing with eSTS patients: being able to objectively counsel patients regarding their personalized residual risk during FU. Patients might be comforted from an improvement in prognosis as the time goes by without events and the update of the prognostic estimate may also support patients' planning for the future. Moreover, the dynamic prediction informs the physician of how a postoperative event will impact on an individual patient's prognosis quantitatively. Finally, this study paves the way for future FU personalization with possible creation of risk-adapted FU strategies. This new nomogram has also been incorporated in the app 'Sarculator' for smartphones and tablets, which is available for free download [35].

Funding

Declaration of Competing Interest

None.

Acknowledgements

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.eclinm.2019.11.008.

References

[1] Fletcher CDM, Bridge JA, Hogendoorn PCW, Mertens F. WHO classification of tumours of soft tissue and bone. Lyon: International Agency for Research on Cancer; 2013.

[2] Callegaro D, Miceli R, Bonvalot S, Ferguson P, Strauss DC, Levy A, et al. Development and external validation of two nomograms to predict overall survival and occurrence of distant metastases in adults after surgical resection of localised soft-tissue sarcomas of the extremities: a retrospective analysis. Lancet Oncol 2016;17(5):671–80.

[3] Callegaro D, Miceli R, Mariani L, Raut CP, Gronchi A. Soft tissue sarcoma nomograms and their incorporation into practice. Cancer 2017;123(15):2802–20.

[4] Mariani L, Miceli R, Kattan MW, Brennan MF, Colecchia M, Fiore M, et al. Validation and adaptation of a nomogram for predicting the survival of patients with extremity soft tissue sarcoma using a three-grade system. Cancer 2005;103(2):402–8.

[5] Cahlon O, Brennan MF, Jia X, Qin LX, Singer S, Alektiar KM. A postoperative nomogram for local recurrence risk in extremity soft tissue sarcomas after limb-sparing surgery without adjuvant radiation. Ann Surg 2012;255(2):343–7.

[6] van Praag VM, Rueten-Budde AJ, Jeys LM, Laitinen MK, Pollock R, Aston W, et al. A prediction model for treatment decisions in high-grade extremity soft-tissue sarcomas: personalised sarcoma care (PERSARC). Eur J Cancer 2017;83:313–23.

[7] Pasquali S, Colombo C, Pizzamiglio S, Verderio P, Callegaro D, Stacchiotti S, et al. High-risk soft tissue sarcomas treated with perioperative chemotherapy: improving prognostic classification in a randomised clinical trial. Eur J Cancer 2018;93:28–36.

[8] Pasquali S, Pizzamiglio S, Touati N, Litiere S, Marreaud S, Kasper B, et al. The impact of chemotherapy on survival of patients with extremity and trunk wall soft tissue sarcoma: revisiting the results of the eortc-stbsg 62931 randomised trial. Eur J Cancer 2019;109:51–60.

[9] Abbott AM, Habermann EB, Parsons HM, Tuttle T, Al-Refaie W. Prognosis for primary retroperitoneal sarcoma survivors: a conditional survival analysis. Cancer 2012;118(13):3321–9.

[10] Parsons HM, Habermann EB, Tuttle TM, Al-Refaie WB. Conditional survival of extremity soft-tissue sarcoma: results beyond the staging system. Cancer 2011;117(5):1055–60.

[11] Miller BJ, Lynch CF, Buckwalter JA. Conditional survival is greater than overall survival at diagnosis in patients with osteosarcoma and ewing's sarcoma. Clin Orthop Relat Res 2013;471(11):3398–404.

[12] Ou JY, Spraker-Perlman H, Dietz AC, Smits-Seemann RR, Kaul S, Kirchhoff AC. Conditional survival of pediatric, adolescent, and young adult soft tissue sarcoma and bone tumor patients. Cancer Epidemiol 2017;50(Pt A):150–7.

[13] Rueten-Budde AJ, van Praag VM, PERSARC studygroup, van de Sande MAJ, Fiocco M. Dynamic prediction of overall survival for patients with high-grade extremity soft tissue sarcoma. Surg Oncol 2018;27(4):695–701.

[14] van Houwelingen HC. Dynamic prediction by landmarking in event history analysis. Scand J Stat 2007;34:70–85.

[15] Fontein DB, Klinten Grand M, Nortier JW, Seynaeve C, Meershoek-Klein Kranenbarg E, Dirix LY, et al. Dynamic prediction in breast cancer: proving feasibility in clinical practice using the team trial. Ann Oncol 2015;26(6):1254–62.

[16] Trojani M, Contesso G, Coindre JM, Rouesse J, Bui NB, de Mascarel A, et al. Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. Int J Cancer 1984;33(1):37–42.

[17] Akaike H. Information theory as an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, editors. Second international symposium on information theory. Budapest, Hungary: Akademiai Kiado; 1973. p. 267–81.

[18] Institute for Statistics and Mathematics of WU (Wirtschaftsuniversität Wien): The R Project for Statistical Computing. http://www.r-project.org/.

[19] Wilson R. The challenge of sarcomas: the patient advocacy group perspective. Clin Sarcoma Res 2019;9 11-019-0121-6. eCollection 2019.

[20] Cupit-Link M, Syrjala KL, Hashmi SK. Damocles' syndrome revisited: update on the fear of cancer recurrence in the complex world of today's treatments and survivorship. Hematol Oncol Stem Cell Ther 2018;11(3):129–34.

[21] Koch L, Jansen L, Brenner H, Arndt V. Fear of recurrence and disease progression in long-term (>/= 5 years) cancer survivors−a systematic review of quantitative studies. Psychooncology 2013;22(1):1–11.

[22] Curda A. The damocles syndrome: where we are today. J Cancer Educ 2011;26(2):397–8.

[23] Toulmonde M, Le Cesne A, Mendiboure J, Blay JY, Piperno-Neumann S, Chevreau C, et al. Long-term recurrence of soft tissue sarcomas: prognostic factors and implications for prolonged follow-up. Cancer 2014;120(19):3003–6.

[24] Barry MJ, Edgman-Levitan S. Shared decision making−pinnacle of patient-centered care. N Engl J Med 2012;366(9):780–1.

[25] Moschetti I, Cinquini M, Lambertini M, Levaggi A, Liberati A. Follow-up strategies for women treated for early breast cancer. Cochrane Database Syst Rev 2016(5):CD001768 doi(5):CD001768.

[26] Mokhles S, Macbeth F, Farewell V, Fiorentino F, Williams NR, Younes RN, et al. Meta-analysis of colorectal cancer follow-up after potentially curative resection. Br J Surg 2016;103(10):1259–68.

[27] Gronchi A, Lo Vullo S, Colombo C, Collini P, Stacchiotti S, Mariani L, et al. Extremity soft tissue sarcoma in a series of patients treated at a single institution: local control directly impacts survival. Ann Surg 2010;251(3):506–11.

[28] Gronchi A, Stacchiotti S, Verderio P, Ferrari S, Martin Broto J, Lopez-Pousa A, et al. Short, full-dose adjuvant chemotherapy (CT) in high-risk adult soft tissue sarcomas (STS): long-term follow-up of a randomized clinical trial from the italian sarcoma group and the spanish sarcoma group. Ann Oncol 2016;27(12):2283–8.

[29] Rydholm A, Gustafson P. Should tumor depth be included in prognostication of soft tissue sarcoma? BMC Cancer 2003;3:17.

[30] Italiano A, Le Cesne A, Mendiboure J, Blay JY, Piperno-Neumann S, Chevreau C, et al. Prognostic factors and impact of adjuvant treatments on local and metastatic relapse of soft-tissue sarcoma patients in the competing risks setting. Cancer 2014;120(21):3361–9.

[31] Yang JC, Chang AE, Baker AR, Sindelar WF, Danforth DN, Topalian SL, et al. Randomized prospective study of the benefit of adjuvant radiation therapy in the treatment of soft tissue sarcomas of the extremity. J Clin Oncol 1998;16(1):197–203.

[32] Callegaro D, Miceli R, Bonvalot S, Ferguson P, Strauss DC, Levy A, et al. Impact of perioperative chemotherapy and radiotherapy in patients with primary extremity soft tissue sarcoma: retrospective analysis across major histological subtypes and major reference centres. Eur J Cancer 2018;105:19–27.

[33] Biau DJ, Ferguson PC, Turcotte RE, Chung P, Isler MH, Riad S, et al. Adverse effect of older age on the recurrence of soft tissue sarcoma of the extremities and trunk. J Clin Oncol 2011;29(30):4029–35.

[34] Wunder JS, Healey JH, Davis AM, Brennan MF. A comparison of staging systems for localized extremity soft tissue sarcoma. Cancer 2000;88(12):2721–30.

[35] DIGITAL FOREST srl. Applestore: https://geo.itunes.apple.com/us/app/sarculator/id1052119173?mt=8 Google Play: https://play.google.com/store/apps/details?id=it.digitalforest.sarculator Nov, 13th 2015. Accessed Oct, 12th, 2019.